LEVEL II

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 19 16669.9-M 18 ARO	2. GOVT ACCESSION NO. AD-A105 907	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Kinds of Polyconfidence Intervals for Centers, and Some Thoughts on Identification and Selection of Confidence Procedures Using Configural Polysampling		5. TYPE OF REPORT & PERIOD COVERED Technical rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 John W. Tukey	12	8. CONTRACT OR GRANT NUMBER(s) 15 DAAG29-79 C 0205
9. PERFORMING ORGANIZATION NAME AND ADDRESS Princeton University Princeton, NY 08544		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 13 34
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709	11	12. REPORT DATE 16 Apr 81
		13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 14 TR-194-SER-2		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC
ELECTE
OCT 2 1 1981
H

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

If we deal initially only with single situations, monoconfidence intervals can be either required only to be weak, to have proper confidence when averaged over all configurations, or to be strong, when they must have the appropriate properties conditionally for each configuration. Assuming location-and-scale invariance throughout both for situations and confidence intervals. When one considers a plot involving several situations, three kinds of polyconfidence intervals seem worth mention: (a) doubly-strong, whose properties hold separately for each

DTIC FILE COPY
AD A105907

20.   ABSTRACT CONTINUED

combination of a situation and a configuration, (b) weak, whose properties hold on average for each situation, (c) singly strong, where the properties hold, first on average for each situation and, second, for each configuration and one situation.   These notions, as well as that of balance in a conservative sense, are explored in the framework of configurations.

Accession For		
NTIS GRA&I	✓	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A		

Kinds of Polyconfidence Intervals for Centers, and
Some Thoughts on Identification and Selection of
Confidence Procedures Using Configural
Polysampling*

by

John W. Tukey
Princeton University
Princeton, New Jersey 08544
and
Bell Laboratories
Murray Hill, New Jersey 07974

Technical Report No. 190, Series 2
Department of Statistics
Princeton University
April 1981

10 20 81

ABSTRACT

If we deal initially only with single situations, monoconfidence intervals can be either required only to be weak, to have proper confidence when averaged over all configurations, or to be strong, when they must have the appropriate properties conditionally for each configuration. (We assume location-and-scale invariance throughout both for situations and confidence intervals.) When we consider a plot involving several situations, three kinds of polyconfidence intervals seem worth mention: (a) doubly-strong, whose properties hold separately for each combination of a situation and a configuration, (b) weak, whose properties hold on average for each situation, (c) singly strong, where the properties hold, first on average for each situation and, second, for each configuration and one situation. These notions, as well as that of balance in a conservative sense, are explored in the framework of configurations.

Attention then shifts to finding such polyconfidence intervals, using configural polysampling* as the principal tool.

*See Technical Report No. 185, by Pregibon and Tukey for general background.

A

INTRODUCTION

1. The one-situation Framework

We begin with the classic case of a single location and
scale situation, whereby

$$y = (y_1, y_2, \ldots, y_n)$$

is, for a simple situation, a set of n iid quantities or,
for a compound situation, a set of n evid quantities, whose
underlying distributions are in either case completely
specified except for location and scale changes.

Here "iid" stands for independently and identically
distributed and "evid" stands for "an exchangeable version
of independently distributed". The latter means that
$y_1, y_2, \ldots, y_n$ are an unknown, equiprobable permutation of
$z_1, z_2, \ldots, z_n$ which are independent with z_1 distributed
according to $f_i(z)$. The one wild Gaussian situation is a
classical example. (Still more complex sorts of situation
have not, as yet, been introduced.)

We find it somewhere between convenient and essential
to work with the order statistic form of the y's, which
could be termed "oiid" in the simple case and "ovid" in the

April 16, 1981

compound case. So we take $y_1 \leq y_2 \leq \cdots \leq y_n$ as certain.

Our concern is with estimating a location μ, which we assume is defined as part of each situation. We require location-and-scale invariance, both of the situation and for the estimate, and thus find it natural to work with location-and-scale configurations, conveniently parameter-ized (for any choice of a and b with $1 \leq a \leq b \leq n$) by

$$y_a = r$$

$$y_b = r+s$$

$$y_i = r+sc_i \qquad \text{for all other i}$$

Here in view of the ordering, we must have

$$c_i < 0 \quad \text{all } i < a$$
$$0 < c_i < 1 \quad \text{all } i \text{ with } a < i < b$$
$$1 < c_i \qquad \text{all } i > b$$

When doing numerical calculations it seems desirable to take a near n/4 and b near 3n/4.

We call the (n-2)-component vector $= c_1, c_2, \ldots(),$ $\ldots(), c_n$, the configuration (more precisely but rarely the (a,b)-location-and-scale-configuration). Here the ()'s remind us, this once only, of the omission of coordinates for $i = a$ and $i = b$.

The configuration, and the two anchors, y_a and y_b, together

April 16, 1981

♦ determine uniquely, and are uniquely determined by the observed y's.

♦ are all that is known to us

♦ are the basis on which we are to construct estimates, confidence intervals, etc. Because of the invariance requirement, we can only use the anchors y_a and y_b linearly. That is, only numbers of the form

$$y_a + t(y_b - y_a)$$

can be used for estimates, confidence interval end points, and the like, where t will usually be a function of the configuration c.

2. Monoconfidence Intervals

In more graphic terms, we have two anchors, y_a and y_b, and a picture of the configuration including o where c_a would otherwise be and 1 where c_b would otherwise be. The configuration is pure shape; the anchors are for location and scale. We look at the picture and decide where in the picture we wish to put a point (an estimate, a confidence interval endpoint, etc.) If this is to be at t in the c-pictures the corresponding y-like value is $y_a + t(y_b - y_a)$. Our requirement of location-and-scale invariance does not allow us to even think about the actual values of y_a and y_b while choosing t.

A pair of functions of configuration L(c), U(c) define

April 16, 1981

an [exact, conservative] p% monoconfidence interval for μ
for a given situation Q, if

$$\text{Prob}_Q\{r + L(c)s \leq \mu \leq r + U(c)s\}$$

is [equal to, greater than or equal to] p/100. Here $\text{Prob}_Q\{\}$
applies equally to each and every set of underlying distri-
butions in the situation Q. (These differ only by
location-and-scale changes. If equality or inequality holds
for one instance of the situation, it holds for all.)

We can easily ask a little more. The ends, $r + sL(c)$
and $r + sU(c)$, of our monocinfidence interval divide $(-\infty,$
$\infty)$ into three intervals. We may also want some sort of
balance. The natural requirement is that neither should
have an excessive chance of containing μ, that both

$$\text{Prob}_Q\{\mu \leq r + sL(c)\} \leq \frac{100-p}{200}$$

and

$$\text{Prob}_Q\{r + sU(c)\} \leq \frac{100-p}{200}$$

When these "end conditions" hold we will speak of a balanced
monoconfidence interval. Such a monoconfidence interval
provides both two-sided and one-sided confidence statements.

We know that there can be many different kinds of mono-
confidence intervals, for a given situation Q. Fisher ( )
introduced the idea of requiring behavior of estimates, etc.
to hold separately for each recognizable distinction among

April 16, 1981

the data possibilities considered. In our set-up, this would mean that our confidence should hold conditionally upon the configuration c, so that

$$\text{Prob}_Q\{r + L(c)s \leq \mu \leq r + U(c)s | c\}$$

is [equal to, greater than or equal to] p/100. We will call a monoconfidence limit fulfilling such a condition a <u>strong</u> monoconfidence interval.

Such a requirement greatly reduces flexibility in choosing confidence intervals. More precisely:

    ◆ it rules out many wasteful choices for monoconfidence intervals, and

    ◆ in special circumstances (e.g. Cox 19..) keeps us from making swaps of probability between recognizably different bodies of data which could lead to apparent overall gain. By restricting ourselves to strong mono-confidence intervals, we are immune to challenges of the form "but look at your configuration, you know what happens under situation Q with such configurations". As a result, challenges will have to be to the situation itself.

* the con-con function *

For any real t, the value of

April 16, 1981

$$\text{Prob}_Q\{\mu \leq t|c\} = G_{Qc}(t)$$

depends only on the indicated arguments, Q and c. As generating a <u>con</u>ditional <u>con</u>fidence statement based on the <u>con</u>figuration, it is mnemonic to call $G_{Qc}()$ the con-con function, given Q and c.

Every strong monoconfidence interval under situation Q arises by satisfying

$$G_{Qc}(L(c)) \leq d$$

$$G_{Qc}(U(u)) \geq d + \frac{p}{100}$$

for a suitable d, as is easy to see.

The minimal (i.e., not purely shortenable) strong monoconfidence intervals, given Q and c arise from equality in these two inequalities. Moreover, there is a minimal balanced strong monoconfidence interval given Q. This is found by solving

$$G_{Qc}(L(c)) = \frac{100-p}{200}$$

$$G_{Qc}(U(c)) = \frac{100+p}{200}$$

3. <u>Qualitative</u> <u>Discussion</u> <u>of</u> <u>Polyconfidence</u> <u>Intervals</u>: <u>Doubly-strong</u> <u>and</u> <u>Weak</u> <u>Instances</u>

Since challenge is now restricted to challenging Q, it is both natural and important to consider at least several Q's. For a qualitative discussion, we do not need to say

April 16, 1981

just how many we consider, so let {Q} be a collection of
situations. Notice, that we may as well call a collection
of situations a plot.

Our concern is then with kinds of polyconfidence situa-
tions as seen in the light of a specified plot.

* doubly-strong polyconfidence *

The strongest requirement we could make is that of a
doubly-strong polyconfidence interval, for which we require

$$\text{Prob}_Q\{r + sL(c) \leq \mu \leq r + sU(c)/c\} \frac{p}{100}$$

for all Q and c. So long as we restrict ourselves to Q's in
a particular plot, we can hardly ask very much more than
this. (Indeed, it may well be that we are asking so much
that we would be willing to take less.)

We can try to ask a little more, however, namely bal-
ance. This requires that, for all c and all Q in the plot
Π, both

$$G_{Qc}(L(c)) \leq \frac{100-p}{200}$$

and

$$G_{Qc}(U(c)) \geq 1 - \frac{100-p}{200} = \frac{100+p}{200}$$

If we define a pair of plot con-cons by

$$G_{-\Pi c} = \min\{G_{Qc}(t) | Q \text{ in } \Pi\}$$

$$G_{+\Pi c} = \max\{G_{Qc}(t) | Q \text{ in } \Pi\}$$

then a balanced doubly strong confidence interval is one for which

$$G_{+\Pi c}(L(c)) \leq \frac{100-p}{200}$$

$$G_{-\Pi c}(U(c)) \geq \frac{100+p}{200}$$

so that the minimal (tightest) such is given by

$$L(c) = G_{+\Pi c}^{-1} \left| \frac{100-p}{200} \right|$$

$$U(c) = G_{-\Pi c}^{-1} \left| \frac{100-p}{200} \right|$$

For a particular p and Π, it is possible that there are no solutions to these equations. This could happen if the con-cons $G_{Qc}(t)$ for individual Q's differed too much. (This can't happen for plots containing a finite number of Q's.)

* weak polyconfidence *

A doubly strong balanced polyconfidence interval is immune to any challenge that does not challenge the plot Π. Accordingly, it will tend to be a long interval. If we ask less, we might shorten it, perhaps considerably on average.

Reducing the size of the plot would surely do this, but suppose this is not desired. We could give up all trace of the "strong" requirement, and ask only that

April 16, 1981

$$\text{Prob}_Q\{r + sL(c) <+ \mu \leq r + sU(c) = \frac{p}{100}$$

for each Q in Π and on the average over its c's. This would
make the interval $(r + sL(c), r + sU(c)$ at least a <u>weak</u>
polyconfidence interval. There are such intervals, the
sign-test-based non-parametric interval for the median being
a balanced one. Note that this particular weak polyconfi-
dence interval will not be strong for the same, nominal
value of p, at least if the pure Gaussian situation is in
the plot. This is so because preserving conditional proba-
bilities for the Gaussian situation requires using limits of
the form

$$\overline{y} + t^* s^*$$

where

$$(s^*)^2 = \frac{1}{n-1} \Sigma(y_i - \overline{y})^2$$

Something quite different from limits of the form

$$y_a + t(y_b - y_a)$$

There are many weak polyconfidence intervals and a dis-
cussion of how to choose one could indeed be lengthy. We
note that such nonparametric procedures as the sign test and
the (one-sample) Wilcoxon test offer very clearly specified
examples.

4. <u>A</u> <u>Compromise</u>: <u>Singly-Strong</u> <u>Polyconfidence</u> <u>Intervals</u>

April 16, 1981

It is natural to believe that:

● doubly-strong polyconfidence intervals are wastefully long,

● weak polyconfidence intervals are too subject to challenge, in terms like "look at that configuration!",

● we need a compromise, where there is at least a reply to such challenges. It is clear what one such compromise would be.

If we knew two things, namely

$$\text{Prob}_Q\{r + sL(c) \leq \mu \leq r + sU(c)\} \geq \frac{p}{100} \quad \text{all } Q \text{ in } \Pi$$

and

$$\text{Prob}_{Q*}\{r + sL(c) \leq \mu \leq r + sU(c)|c\} \geq \frac{p}{100} \quad \text{all } c, \text{ one } Q*$$

Then the answer to "but look at your c" could be "in Q* that wouldn't matter". This deflects the challenge from the knowable configuration to the unknowable situation. For some this would be good enough; for others not. (The latter would have to move to or toward a doubly-strong polyconfidence interval.)

* criss-crossing *

If $L_1(c), U_1(c)$ defines a strong (exact or conservative) monoconfidence interval for $Q*$ (in Π), and $L_2(c), U_2(c)$ defines a weak (exact conservative) polyconfidence interval

April 16, 1981

over Π, then $L(c)$, $U(c)$, where

$$L(c) = \min\{L_1(c), L_2(c)\}$$

$$U(c) = \max\{U_1(c), U_2(c)\}$$

will be a singly-strong (conservative) polyconfidence interval since

$$\text{Prob}_{Q*}\{r + sL(c) \leq \mu \leq r + sU(c)|c\}$$

$$\geq \text{Prob}_{Q*}\{r + sL_1(c) \leq \mu \leq r + sU_1(c)|c\} \geq \frac{p}{100}$$

for all c and

$$\text{Prob}_Q\{r + sL(c) \leq \mu \leq r + sU(c)\}$$

$$\geq \text{Prob}_Q\{r + sL_2(c) \leq \mu \leq r + sU_2(c)\} \geq \frac{p}{100}$$

for all Q in Π.

$$* \text{ curtailment } *$$

Even if $(L_1(c), U_1(c))$ is not the same as $(L_2(c), U_2(c))$, we may have

$$\text{Prob}_Q*\{r + sL_2(c) \leq \mu \leq r + sU_2(c)\} \geq \frac{p}{100}$$

for some c's. For such c's we can surely take

$$L(c) = L_2(c)$$

$$U(c) = U_2(c)$$

reserving the "min" and "max" operations for where they are

really needed. Such a curtailed criss-cross is easily
implemented, once we are able to evaluate the probability
above, which equals

$$G_{Q*c}(U_2(c)) - G_{Q*c}(L_2(c))$$

Now we have started economizing, we can continue.

* patching *

How can we tighten our polyconfidence interval further?
One easy way starts with $L_2(c)$ and $U_2(c)$, and looks at

$$G_{Q*c}(L_2(c)) \qquad \text{and} \qquad G_{Q*c}(U_2(c))$$

If these differ by at least $p/100$, we are content with

$$L(c) = L_2(c)$$

$$U(c) = U_2(c)$$

for all such c, and turn our attention to other c's, where

$$G_{Q*c}(U_2(c)) - G_{Q*c}(L_2(c)) < \frac{p}{100}$$

We now look at some one of these other c's, and either
decrease $L_2(c)$ or increase $U_2(c)$ until either starts (on L
or U) showing possible changed values

$$G_{Q*c}(U_2^*(c)) - G_{Q*c}(L_2^*(c)) = \frac{p}{100}$$

or

$$1 - G_{Q*c}(U_2^*(c)) = G_{Q*c}(L_2^*(c))$$

If the former happens first, we stop there, else we continue, preserving the last equality until we have

$$G_{Q*c}(L_2^*(c)) = \frac{100-p}{200}$$

$$G_{Q*c}(U_2^*(c)) = \frac{100+p}{200}$$

Then, relabelling $L_2^1(c)$ as $L(c)$, and $U_2^1(c)$ as $U(c)$ we have the desired values of $L(c)$, $U(c)$ for that particular c.

* tuning *

We are not prepared to offer any meaningful comments about the optimality of such a $L(c)$, $U(c)$ pair as a singly-strong polyconfidence interval. There is a very real possibility that we can do better, but we shall avoid this tuning problem.

* examples *

Let us take n=10, μ = population median, p = 1002/1024~98%, and Π = all reasonable simple situations.

With our conventions,

$$(y_1,\ y_n) = \left| r - \frac{y_a - y_1}{y_b - y_a}\ s,\ r + \frac{y_n - y_b}{y_b - y_a}\ s \right|$$

where

April 16, 1981

$$L_2(c) = -\frac{y_a-y_1}{y_b-y_a} = \frac{y_1-y_1}{y_b-y_a}$$

$$U_2(c) = \frac{y_n-y_b}{y_b-y_a}$$

is a 1002/1024 exact weak polyconfidence interval over Π. Similarly, for $Q^* =$ Gaussian,

$$(\bar{y} - t^*s^*, \; \bar{y} + t^*s^*)$$

where

$$\bar{y} = \frac{1}{n} \Sigma \; y_{\bar{c}}$$

$$(s^*)^2 = \frac{1}{n} \frac{1}{n-1} \Sigma \; (y_i - \bar{y})^2$$

and t^* is the upper 11/1024 percent point of students t on 19 degrees of freedom ($t^* = 2.495$), is a strong monoconfidence interval at Q^*.

The criss-cross $L(c)$, $U(c)$, easily calculated for any configuration, will be a singly-strong polyconfidence interval over all of Π. Almost certainly such a choice will be quite wasteful, since $(s^*)^2$ is easily enlarged by the extreme values of y_1 or y_n. This will be only slightly less true if we curtail the criss-cross, or, probably if we replace it by a patching.

It might well be reasonable to study the results of

◆ using a sign-test or Wilcoxon strong polyconfidence intervals for the median as $L_2(c)$, $U_2(c)$

April 16, 1981

♦ taking Q^* as an intermediate situation, say a simple slacu,

♦ taking $L_1(c)$, $L_1(c)$ as the strong monoconfidence interval on Q^* based on the w5-biweight;

♦ considering all three of criss-cross, curtailment, and patching.

5. The Position

We thus have at least three qualities of polyconfidence intervals,

♦ doubly-strong (easy to find; "wasteful" to some)

♦ strong (allows an answer to the c-challenge, examples easy, tuning not likely to be easy)

♦ weak (many alternatives, tuning not trivial, but probably feasible)

each of which may or may not further include the requirement of balance. If we had examples of all of these for one or more plots, each user could take his/her choice.

Thus we ought to turn to the question of finding such intervals.

April 16, 1981

B

THE CALCULATIONS

5. The Set-up

We look toward configural polysampling as the basis for our practical work. We should have them:

♠ a small plot, consisting of perhaps 2 to 7 situations S_1, S_2, ..., $S_{q\ max}$ for each of which a center, μ, is specified.

♠ a polysample of configurations $\{c\} = c(1)$, ..., $c(m)$ (note that each $c(i)$ is either a $(n-2)$-entry vector array or a n-entry vector array with two fixed entries)

♠ weights W_{jq} appropriate for use when $c(j)$ is to be used as a configuration sampled for S_q.

♠ con-con functions, $G_{qc}(t)$ applicable for these q's and c's.

The character and building of all these pieces, except the con-con function, is discussed in Technical Report 185 (Pregibon and Tukey, 1981) and Technical Report 191 (Bell and Pregibon, 1981) (see also Rogers and Relles, 1973 for formulas in the case $a=1$, $n=b$). The basic results depend on averages, for fixed c, represented first as integrations over r and s and then on integration over a rectangle (where Gaussian quadrature formulas apply each way). If $I_{Qc}(r,s,t)$ is the indicator function

April 16, 1981

$$I_{Qc}(r, st) = 1, \text{ if } \mu \le y_a + t(y_b - y_a) = r + st$$

$$= 0, \text{ else}$$

which depends on q (where $Q = S_q$) and c then

$$G_{Qc}(t) = \text{Prob}_Q\{\mu \le r + st\} = \text{ave}_{Qc}\{I(r, s_j t)\}$$

and this can be evaluated for selected values of t in a way similar to the other integral evaluations.

7. Doubly-strong Polyconfidence Intervals

To find the monimal balanced 95% polyconfidence interval, for each of the configurations of the polysample, we have merely to:

♠ evaluate the con-con functions $G_{qc}(t)$ for each q in the plot, each c in the polysample, and well-selected values of t,

♠ calculate the + and - con-con functions from

$$G_{-\Pi c}(t) = \min (G_{qc}(t)/S_q \text{ in } \Pi)$$

$$G_{+\Pi c}(t) = \max (G_{qc}(t)/S_q \text{ in } \Pi)$$

the former for higher values of t, the latter for lower values of t,

♠ solve the equations

$$G_{+\Pi c} = 2.5\%$$

$$G_{-\Pi c}(t) = 97.5\%$$

for each c in the polysample.

♦ take these values of t as $L(c)$ and $U(c)$

♦ assert, <u>for these configurations</u> (and any others to be treated later), that the interval in question is from

$$r + sL(c) = y_a + L(c)(y_b - y_a)$$

to

$$r + sU(c) = y_a + U(c)(y_b - y_a)$$

If we want to know how this polyconfidence intervals performs, we average what we see at the given polysamples using the appropriate weights. Thus the average lengths of our confidence intervals, which we do <u>NOT</u> think is likely to be a good criterion to consider would be found as an estimate of

$$\text{ave}_{qc} \ \text{ave}_{r,s} \ (U(c) - L(c))(s)$$

$$\text{ave}_{(q)c}(\text{ave}_{(qc)r,s}(U(c)-L(c))(s)) = \text{ave}_{(q)c}(U(c)-L(c)) \ \text{ave}_{(qc)}s^s)$$

perhaps as

$$\frac{1}{\sum\limits_{c} W_{qc}} \sum_{c} W_{qc}(U(c) - L(c))\bar{s}_{qc}$$

where

$$\bar{s}_{qc} = \text{ave}_{(qc)}\{s\}$$

was itself estimated by Gaussian quadrature.

Notice that there are q-max different such average lengths.

April 16, 1981

Thus we need not expect too much difficulty (a) in finding L(c) and U(c) and (b) in evaluating simple properties.

8. Tuning Weak Monoconfidence Intervals for Average Lengths

Before we attempt to tune more complicated structures, it is well to consider tuning weak monoconfidence intervals using the same dubious criterion of average length. Here we wish to choose $L(c)$ and $U(c)$ to satisfy

$$\underset{c}{\overset{Q}{\text{ave}}} \{G_{Qc}(U(c)) - G_{Qc}(L(c))\} \geq \frac{p}{100}$$

while making

$$\underset{c}{\overset{Q}{\text{ave}}} \{(U(c) - L(c))\overline{s}_{QC}\} = \underset{c}{\overset{Q}{\text{ave}}} \{U(c)\overline{s}_{QC}\} - \underset{c}{\overset{Q}{\text{ave}}} \{L(c)\overline{s}_{QC}\}$$

small. (Note that "$\overset{Q}{\text{ave}}$" means what might also be written "ave_q" indicating averaging over the indicated part of a selected instance of the situation; averaging that would include explicit use of weights, were this necessary, -- as it will be in the polysampling case.) If we have, say, 500 configurations on which we are working, we have a constrained optimum problem with 1000 variables. Direct approaches are likely to be inefficient.

We will ordinarily find $G_{Qc}(t)$ ogive-shaped, and its derivative, $g_{Qc}(t)$, unimodal. We will shortly have occasion to be concerned with two inverses of $g_{Qc}(t)$ which we can

define, in general, by

$h_{Qc}^{-}(u)$: the algebraically smallest t with $g_{Qc}(t) = u$,

$h_{Qc}^{+}(u)$: the algebraically largest t with $g_{Qc}(t) = u$,

So long at least as $g_{Qc(t)}$ is continuous, these will be well-defined on $0 \leq u \leq$ max $\{g_{Qc}\}$ (though they might be discontinuous, since $g_{Qc}()$ might not be unimodal). If we write

$$\Sigma_1 = \Sigma b_{Qc}(U(c))$$

$$\Sigma_2 = \Sigma b_{Qc} G_{Qc}(L(c))$$

$$\Sigma_3 = \Sigma b_{Qc} \overline{s}_{Qc} U(c)$$

$$\Sigma_4 = \Sigma b_{Qc} \overline{s}_{Qc} L(c)$$

where the b_{QC} incorporate the needed weights, if any, for the 4 sums that we will use to replace the averages we used to state the problem, we will want to

<u>minimize</u> $\Sigma_3 - \Sigma_4$

subject <u>to</u> $\Sigma_1 - \Sigma_2 \geq \frac{p}{100}$

which is naturally attacked with a Lagrange multiplier, x, by minimizing

$$\Sigma_3 - \Sigma_4 - \lambda(\Sigma_1 - \Sigma_2)$$

and then choosing λ to satisfy the constraint. Since Σ_3 and Σ_1 are functions of the $\{U(c)\}$ above -- and Σ_2 and Σ_4 of the

$\{L(c)\}$ -- we can extremelize

$$\Sigma_3 - \lambda\Sigma_1 \qquad \text{and} \qquad \Sigma_4 - \lambda\Sigma_2$$

separately (unless this leads to some $U(c) <$ the corresponding $L(c)$.) This leads to differentiating wirt. $U(c)$, to

$$b_{Qc}\bar{s}_{Qc} - \lambda b_{Qc}g_{Qc}(U(c)) = 0$$

whence we may take

$$U(c) = h^+_{Qc}(\bar{s}_{Qc}/\lambda)$$

similarly, we may take

$$L(c) = h^-_{QC}(\bar{s}_{Qc}/\lambda)$$

what remains is to empirically choose λ to make

$$\operatorname*{ave}_{c}{}^{Q} \{G_{Qc}(U(c)) - G_{QC}(L(c))\}$$

equal to the desired p/100.

9. Other Criteria

We also want to be able to tune our monoconfidence intervals for other criteria, which deserve some discussion in their own right.

We all recognize some form of confidence interval as the best we can do. And few of us want a 99.9999% interval. Thus we have come to accept a meaningful but small (say 5% or 1%) chance that our confidence interval will not cover

the center at which it is aimed. Should we then pay too
much attention to a small chance that the interval is very
long?

When the interval is too long, it is unhelpful, but we
know it. Can this be nearly as important as missing its
target? It would seem that the answer should be a rousing
"no"! If we are going to allow 5% misses, then we ought, we
very well argue, accept 10% overlong intervals. So we seek
related criteria.

Two natural, but possibly naive, choices are:

◆ the 90% point of the length of the confidence inter-
vals, and

◆ the average length of the shortest 90% of all confi-
dence intervals.

If we are doing direct sampling, either of these can be used
simply and directly, just by sorting the empirical interval
lengths.

If we are working with configurations, we have to add a
loop. For what can be reasonably calculated at a configura-
tion is the chance that an interval -- cr a configurations
-- should be shorter than a prescribed length. If we
prescribe a length, moderately extensive computation gives
us a % less than this length. Then we have to adjust the
length, and iterate. If we must, we must.

April 16, 1981

But let us think about our criteria with a little more care. Perhaps the last two assess, separately, two aspects that we might like to assess together. Suppose that we took

$$\frac{9}{10} \text{ ave \{90\% shortest lengths\}} + \frac{1}{10} \text{ (90\% point)}$$

This is the average of a saturating function

= length,   below the bend at the 90% length

= 90% length,   else

There are possible virtues to such a combination. Let K be a trial value for the 90% length, then we would estimate

ave {length|length $\leq$ K}

Prob {length $\leq$ K}

which is equivalent to estimating

ave {K - length|length $\leq$ K}

Prob {length $\leq$ K}

where we still must iterate on K to make the probability = to p/100. The criterion will then take the value

$$K - \text{ave}\{K - \text{length}|\text{length} \leq K\} = K - K \text{ ave}\{1 - \frac{\text{length}}{K}|\text{length} \leq K\}$$

where the last factor is a relatively slowly changing function of K.

10. Iterating for the Select Criterion

Let us put

$$A_{Qc}(u) = \operatorname*{ave}_{s} \left| (1 - \frac{s}{u}) \right| \text{ given } s \leq u \Big|$$

$$B_{Qc}(u) = \operatorname*{prob}_{r,s} \{s \leq u\}$$

so that the probability that the interval length be less than K is, for one c,

$$B_{Qc} \left| \frac{K}{U(c) - L(c)} \right|$$

so that we must eventually control

$$\Sigma b_{Qc} \; B_{Qc} \left| \frac{K}{U(c) - L(c)} \right|$$

by changing K.   For K fixed, however, we desire to <u>minimize</u> K-(K times the following) and hence, for fixed K, to <u>maximize</u>

$$\Sigma b_{Qc} \; A_{Qc} \left| \frac{K}{U(c) - L(c)} \right|$$

since

$$1 - \frac{\text{length}}{K} = 1 - \frac{(U(c) - L(c))s}{K} = 1 - \frac{s}{K/(U(c) - L(c))}$$

The Lagrangian form to be extremal is now

$$\Sigma b_{Qc} \; A_{Qc} \left| \frac{K}{U(c) - L(c)} \right| - \lambda \; (\Sigma b_{Qc} G_{QC}(U(c)) - \Sigma b_{Qc} G_{QC}(L(c))$$

whose derivatives wirt U(c), and L(c) with

$$a_{Qc}(u) = \frac{\partial}{\partial u} a_{Qc}(u)$$

are, less the common factor b_{Qc}

April 16, 1981

$$- a_{Qc} \left| \frac{K}{U(c)-L(c)} \right| \cdot \frac{K}{(U(c)-L(c))^2} - \lambda g_{Qc}(U(c)),$$

and

$$+ a_{Qc} \left| \frac{K}{U(c)-L(c)} \right| \cdot \frac{K}{(U(c)-L(c))^2} - \lambda g_{Qc}(L(c))$$

If these vanish, so does their difference, hence

$$g_{Qc}(U(c)) = g_{Qc}(L(c))$$

as tacitly before, giving

$$U(c) = b_{Qc}^+(g_{Qc}(L(c)))$$

and

$$U(c) = L(c) = A_{Qc}^+(g_{Qc}(g_{Qc}(L(c))) - L(c) = f_{QC}(L(c))$$

so that

$$\lambda g_{Qc}(L(c)) = a_{Qc} \left| \frac{K}{f_{Qc}(L(c))} \right| \frac{K}{(f_{Qc}(L(c))^2}$$

which should be soluble for $L(c)$, given λ possibly with some effort. Once this is done for all c in our sample of configurations, we will again want to check

$$\text{Prob} \{\text{length} \leq K\} = \Sigma b_{Qc} \ B_{Qc} \left| \frac{K}{U(c) - L(c)} \right|$$

and adjust K to bring this to the desired value.

The process is appreciably more complicated than for the average length criterion, but apparently not unbearably so.

April 16, 1981

11. Tuning Weak Polyconfidence Intervals

Suppose we want a weak polyconfidence interval for a plot consisting of two situations, A and Z. We now have, if we stick to the simple case of the average length criterion, eight Σ's to consider:

$$\Sigma_{A1} = \Sigma b_{Ac} G_{Ac}(U(c))$$

$$\Sigma_{A2} = \Sigma b_{Ac} G_{Ac}(L(c))$$

$$\Sigma_{Z1} = \Sigma b_{Zc} G_{Zc}(U(c))$$

$$\Sigma_{Z2} = \Sigma b_{Ac}(L(c())$$

$$\Sigma_{A3} = \Sigma b_{Ac} s_{\overline{Ac}} U(c)$$

$$\Sigma_{A4} = \Sigma b_{Ac} s_{\overline{Ac}} L(c)$$

$$\Sigma_{Z3} = \Sigma b_{Zc} s_{\overline{Zc}} U(c)$$

$$\Sigma_{Z4} = \Sigma b_{Zc} s_{\overline{Zc}} L(c)$$

and we desire to minimize, jointly,

$$\Sigma_{A3} - \Sigma_{A4}$$

and

$$\Sigma_{Z3} - \Sigma_{Z4}$$

subject to

April 16, 1981

$$\Sigma_{A1} - \Sigma_{A2} \geq p/100$$

$$\Sigma_{Z1} - \Sigma_{Z2} \geq p/100$$

This clearly calls for one pair (d,ξ) of shadow prices and one pair of Lagrange multipliers, all of which leave us minimizing

$$\mathsf{d}(\Sigma_{A3} - \Sigma_{A4}) + \xi(\Sigma_{Z3} - z_{Z4}) - \lambda_A(\Sigma_{A1} = \Sigma_{A2}) - \lambda_Z(\Sigma_{Z1} - \Sigma_{Z2})$$

which can again be done separately for two parts, here minimizing

$$\mathsf{d}\Sigma_{A3} + \xi\Sigma_{Z3} - -\lambda_A\Sigma_{A1} - \lambda_Z\Sigma_{Z1}$$

and maximizing

$$\mathsf{d}\Sigma_{A4} + \xi\Sigma_{Z4} - \lambda_A\Sigma_{A2} - \lambda_Z\Sigma_{Z2}$$

which lead, on differentiating w.r.t. $U(c)$ and $L(c)$, respectively, to

$$0 = \mathsf{d}b_{Ac}\bar{s}_{Ac} + \xi b_{Zc}\bar{s}_{Ac} - \lambda_A b_{Ac}g_{Ac}(L(c)) - \lambda_Z b_{Zc}g_{Zc}(U(c))$$

and

$$0 = \mathsf{d}b_{Ac}\bar{s}_{Ac} + \xi b_{Zc}\bar{s}_{Ac} - \lambda_A b_{Ac}g_{Ac}(L(c)) - \lambda_Z b_{Zc}g_{Zc}(L(c))$$

If we now write

$$\lambda_A = \lambda\theta$$

$$\lambda_Z = \lambda(1-\theta)$$

April 16, 1981

both right-hand pairs of terms can be written in terms of

$$\theta b_{Ac} g_{Ac}(t) + (1-\theta) b_{Zc} g_{Bc}(t) = h_{\theta c}^{-1}(t)$$

whose inverses can be written as

$$h_{\theta c}^{+}(u) \quad \text{and} \quad h_{\theta c}^{-}(u)$$

so that we have

$$0 = \alpha b_{Ac} \bar{s}_{Ac} + \xi b_{Zc} \bar{s}_{Zc} - \lambda h_{\theta c}^{-1}(U(c))$$

$$0 = \alpha b_{Ac} \bar{s}_{Ac} + \xi b_{Zc} \bar{s}_{Zc} - \lambda h_{\theta c}^{-1}(L(c))$$

and

$$U(c) = h_{\theta c}^{+} \left| \frac{\alpha b_{Ac} \bar{s}_{Ac} + \xi b_{Zc} \bar{s}_{Zc}}{\lambda} \right|$$

$$L(c) = h_{\theta c}^{-} \left| \frac{\alpha b_{Ac} \bar{s}_{Ac} + \xi b_{Zc} \bar{s}_{Zc}}{\lambda} \right|$$

where θ and λ must now be varied jointly to ensure

$$\Sigma b_{Ac} G_{Ac}(U(c)) - \Sigma b_{Ac} G_{Ac}(L(c)) \geq p/100$$

$$\Sigma b_{Zc} G_{Zc}(U(c)) - \Sigma b_{Zc} G_{Zc}(L(c)) \geq p/100$$

Outside of this loop, we must vary α/ξ (we will simplify matters by forcing, say, $\alpha + \xi = 2$) in order to get the right joint minimum for the two average lengths.

Plausibly what we may seek at this point is

April 16, 1981

$$\frac{\overset{A}{\text{ave}}\{\text{polyconfidence length}\}}{\underset{A}{\text{min}}\ \overset{A}{\text{ave}}\{\text{monoconfidence length}\}} = \frac{\overset{Z}{\text{ave}}\{\text{polyconfidence lengths}\}}{\underset{Z}{\text{min}}\ \overset{Z}{\text{ave}}\{\text{monoconfidence length}\}}$$

thus maximizing a polyefficiency defined by

polyefficiency = min{monoefficiencies}

$$A - \text{monoefficiency} = \left|\frac{A-\text{minimum average length}}{\text{actual minimum length}}\right|2$$

$$Z - \text{monoefficiency} = \left|\frac{Z-\text{minimum average length}}{\text{actual minimum length}}\right|2$$

12. Comment

The calculations are clearly getting moderately complex. They will get somewhat worse for other criteria or more situations. But they seem likely to be feasible.

13. Tuning Singly-strong Confidence Intervals

If we have also fixed a situation Q^*, at which we want our polyconfidence internal to be a strong monoconfidence interval, what we have done is to require

$$G_{Qc}(U(c)) - G_{Qc}(L(c)) \geq p/100$$

for each c.

We expect the usual situation. For some c (given, say, d/ξ, λ, and θ) this condition will be satisfied, for others not. For the others we must involve Q in the choice of U(c) and L(c).

We can easily parametrize the possibilities in terms of L(c). Given a sufficiently small L(c) (for given $d/\xi, \lambda$, and θ), there will be:

   • the smallest U(c) that satisfies the condition displayed above,

   • a minimum cost U(c), depending upon $d/\xi, \lambda$, and theta.

consider the larger of these, and its total cost, at c of the (L(c),U(c)) pair, as given in terms of d, ξ, λ, and θ. Now choose L(c) to minimize this cost and U(c) to be the corresponding larger value.

When this has been done for all c, we are ready to vary d, ξ, λ and θ to obtain the desired result.

Again the process has become somewhat more complicated, but is probably still feasible.

DATE

ILME

10. Iterating for the Select Criterion

Let us put

April 16, 1981